

基于芭蕉属 EST 序列的地涌金莲 SSR 引物开发

李文娟, 马 宏, 李正红*, 万友名, 刘秀贤, 周翠丽

(中国林业科学研究院资源昆虫研究所, 云南 昆明 650224)

摘要:对 31 308 条来自 NCBI 的芭蕉属 (*Musa*) EST 序列进行拼接得到全长为 13.51 Mb 的 21 129 条无冗余 EST 序列 (含 3 818 条 contigs 及 17 311 条 singletons), 其中, 4 944 条 (23.40%) EST 序列含有 5 416 条 SSRs, SSR 的出现频率为 25.63%, 平均分布距离 2.49 Kb, 有 234 条 (1.11%) 含有 1 个以上的 SSR。在所检测的 SSR 中, 二、三、四核苷酸重复是主导重复类型, 分别占 EST-SSR 总数的 21.80% (1 181 条)、52.55% (2 846 条) 和 14.55% (788 条)。AG/CT、AAG/CTT 与 AGG/CCT 和 AAAG/CTTT 与 AAAT/ATTT 分别是二、三、四核苷酸的优势重复基元。随机设计了 238 对 EST-SSR 引物在 24 个地涌金莲个体中进行筛选, 116 对有扩增产物, 其中, 78 对 EST-SSR 引物扩增出清晰稳定的目的片段, 49 对引物表现出多态性。本研究检测的 15 对引物扩增等位基因数范围是 2~7 个, 平均 3.067 个; 表观杂合度 (*Ho*) 范围是 0.042~0.750, 平均 0.250; 期望杂合度 (*He*) 范围是 0.232~0.823, 平均 0.522。

关键词:地涌金莲; EST-SSR; 多态性引物; 分子标记

中图分类号: S718.46

文献标识码: A

Development of SSR Markers of *Musella lasiocarpa* by Data Mining in *Musa* EST Sequences

LI Wen-juan, MA Hong, LI Zheng-hong, WAN You-ming, LIU Xiu-xian, ZHOU Cui-li

(Research Institute of Resources Insects, Chinese Academy of Forestry, Kunming 650224, Yunnan, China)

Abstract: 21 129 non-redundant clusters, containing 3 818 contigs and 17 311 singletons, were identified from a total of 31 308 publicly available *Musa* EST sequences. 4 944 (23.40%) of them contained 5 416 (25.63%) SSR motifs, and the di-(21.80%), tri-(52.55%) and tetra-nucleotide (14.55%) are the main motifs of all SSRs obtained. AG/CT repeats were dominant in di-nucleotide motif, AAG/CTT and AGG/CCT repeats were dominant in tri-nucleotide motif, AAAG/CTTT and AAAT/ATTT repeats were dominant in tetra-nucleotide motif. 238 EST sequences were randomly selected on molecular markers development, for PCR amplification and polymorphism analysis in *Musella lasiocarpa*. Of which, 116 pairs of SSR primers successfully amplified PCR product and 78 pairs gave clear bands. 49 of them were found to be polymorphic. A set of 15 polymorphic SSR markers from above 49 SSR markers selected were analyzed using 24 individuals from 4 wild *M. lasiocarpa* populations. The average allele number of 3.067 per locus was detected with a range from 2 to 7. The observed heterozygosities per marker were ranged from 0.042 to 0.750 (mean 0.250) and expected heterozygosities were 0.232 to 0.823 (mean 0.522).

Key words: *Musella*; EST-SSR; polymorphic primer; molecular markers

微卫星标记 (simple sequence repeat, SSR) 以其特异性好、多态性高、共显性标记、分布广泛和操作简便等优点而迅速成为分子生物学研究的重要手段,但传统的 SSR 引物开发的周期长、费用高且效率低^[1-2]。表达序列标签 (expressed sequence tags, ESTs) 是通过大规模 cDNA 随机测序产生,1 个 EST 代表生物个体的某种组织或细胞在某一时期的 1 个表达基因。利用 EST 开发 SSR 引物相对简单快速、成本低,且在不同种、不同属甚至在不同科的物种间具有良好的通用性^[3-5],同时,急剧上升的 EST 数目为 SSR 的开发提供了可能。1991 年各公共数据库中的 EST 数目不足 2 000 条,而截止到 2011 年 11 月,NCBI 上小麦 (*Triticum aestivum* L.) 有 107 万条,玉米 (*Zea mays* L.) 202 万条,水稻 (*Oryza sativa* L.) 126 万条,芭蕉科 (Musaceae) 植物的 EST 序列超过 3 万条,且仍在不断更新。EST 反映了基因的编码部分,可以直接获得基因表达的信息,与基因组 SSR 相比,EST-SSR 可提供基因表达信息,为功能基因的直接鉴定提供了可能性。目前,EST-SSRs 被广泛应用于遗传连锁图谱的构建、群体遗传学研究、分子辅助育种、品种鉴定和系谱分析^[6-9]等方面的研究。

地涌金莲 (*Musella lasiocarpa* (Franch.) C. Y. Wu ex H. W. Li) 属芭蕉科地涌金莲属 (*Musella*), 是中国特有单属种,多年生大型草本植物^[10],主要分布在我国云南、四川等地,其除具有极高的观赏价值外,可食用、入药^[11]。目前,涉及地涌金莲分子标记的研究很少见,潘庆杰利用 10 对 RAPD 引物对地涌金莲栽培和野生种群进行了遗传多样性分析^[12]; Yang 利用传统测序的方法开发出 17 对地涌金莲 SSR 引物^[13],这远不能满足对其从分子遗传层面进行研究所需。随着新的地涌金莲变异类型和野生种群被发现^[11],对其种群遗传结构的研究迫切需要更多的分子标记;然迄今为止,在公共数据库中还未能搜索到地涌金莲 EST 序列,本文利用生物信息学手段,对数据库中芭蕉属 EST 序列进行分析并开发适用于地涌金莲的 SSR 引物,为后续地涌金莲遗传连锁图谱的构建、遗传多样性的研究奠定基础。

1 材料与方法

1.1 植物材料

选取 4 个地涌金莲野生种群的共 24 个个体,全部表现为绿色叶中脉及纯黄色苞片。取未展开的幼

叶,用 CTAB 法^[14]提取总 DNA,浓度调至 50 ng · μL⁻¹,于 -20 °C 中保存备用。

1.2 EST-SSR 信息分析

2010 年 9 月 10 日从 NCBI (<http://www.ncbi.nlm.nih.gov/nucest>) 以 FASTA 格式获取 31 308 条芭蕉属 EST 序列,主要来自于叶片及根部组织,其中,11 155 条属 *paradisiaca* (11 070 条为 ABB 类型),5 289 为 *M. balbisiana*,14 864 为 *M. acuminata* (7051 为 AAA 类型)。用 sarafer1.3 (<http://acephpx.cropdb.org/serafer.php>) 软件包的 CAP3^[15] 剔除冗余序列,重叠一致百分比 95%,最小序列长度 20 bp,其余参数默认;SeqClean 去除长度小于 150 bp 的序列和劣质序列;LeapSSR 程序搜寻 SSR,SSR 搜寻条件为: di-, tri-, tetra-, penta- 和 hexa-nucleotide (DNRs, TNRs, TeNRs, PNRs 和 HNRs) 最小重复单元分别为 6,4,3,3 和 3,同时搜索复合 SSR (间隔序列不大于 20 bp)。

1.3 引物设计

利用 Primer3^[16] 对含有 SSR 位点的 EST 序列进行引物设计,设计参数:引物长度 18 ~ 25 bp (最适 20 bp);引物 Tm 值 57 °C (最大 63 °C,最小 52 °C,上下游引物 Tm 值相差不超过 5 °C);预期产物大小 100 ~ 500 bp,CG% 含量为 35% 至 65%。引物由上海生工生物工程技术有限公司合成。

1.4 PCR 扩增程序

PCR 反应体系 (20 μL): 25 ng 模板 DNA,0.75U Taq DNA polymerase,1 × PCR buffer (10 mmol Tris-HCl pH 8.0,50 mmol KCl),1.8 mmol MgCl₂,0.13 mmol dNTPs (Takara),上下游引物各 0.25 μmol,去离子水补足 20 μL。PCR 反应在 PTC thermal cycler (BIO-RAD) 上进行:95 °C 预变性 4 min,95 °C 变性 45 s,引物以各自的 Tm 值复性 45 s,72 °C 延伸 1 min,共 30 个循环,最后 72 °C 延伸 10 min。

1.5 检测及数据分析

PCR 产物用 8% (29:1 acclamed: bis) 非变性聚丙烯酰胺凝胶进行分离,电压设定 120 V,时间 140 min,在 15 mg · mL⁻¹ 的 EB (ethidium bromide) 中染色 3 ~ 5 min,凝胶成像仪拍照。产物片段大小用 pUC19 DNA/MspI marker (Fermentas) 进行标定。

以 PopGen32 (version 1.31, <http://www.ualberta.ca/~fyeh/>) 进行等位基因数 (NA)、Linkage Disequilibrium (LD)、表观杂合度 (Observed heterozygosity, *Ho*) 和期望杂合度 (Expected heterozygosity,

He)的分析以及 Hardy-Weinberg equilibrium (HWE) 的计算。

2 结果与分析

2.1 芭蕉属 EST-SSR 发生频率

对 NCBI 上公开的 31 308 条芭蕉属 EST 序列进行分析,去除劣质序列并进行拼接,得到 21 129 条非冗余序列(包括 3 818 条 contigs 和 17 311 条 singletons),序列总长度 13.51 Mb,其中,4 944 条(23.40%)EST 序列含有 5 416 条 SSR,SSR 的出现频率为 25.63%,平均每隔 2.49 kb 就有一个 SSR,有 234 条(1.11%)含有 1 个以上的 SSR。在所检测的 SSR 中,三核苷酸所占比例最多,为 52.55% (2 846 条),其次为二核苷酸,占 21.80% (1 181 条),四核苷酸占 14.55% (788 条),三者共占全部 SSR 的 88.90%,其余五核苷酸(3.86%)和六核苷酸(7.24%)数量相近(表 1)。

SSRs 长度在 12~83 个核苷酸之间,重复次数范围为 3 至 38,绝大多数的 SSR (73.06%)为 3、4、5 或 6 个重复单元(图 1)。

表 1 SSR 在芭蕉属 EST 中的分布

重复类型	数目	所占比例 /%	出现频率 /% ¹⁾	平均距离 /kb ²⁾
二核苷酸	1 181	21.80	5.59	11.44
三核苷酸	2 846	52.55	13.47	4.75
四核苷酸	788	14.55	3.73	17.14
五核苷酸	209	3.86	1.00	64.64
六核苷酸	392	7.24	1.86	34.46
总计	5 416	100.00	25.63	2.49

注:1)出现频率=检出 SSR 数目/无冗余 EST 总数;2)平均距离=无冗余 EST 总长度/SSR 数目。

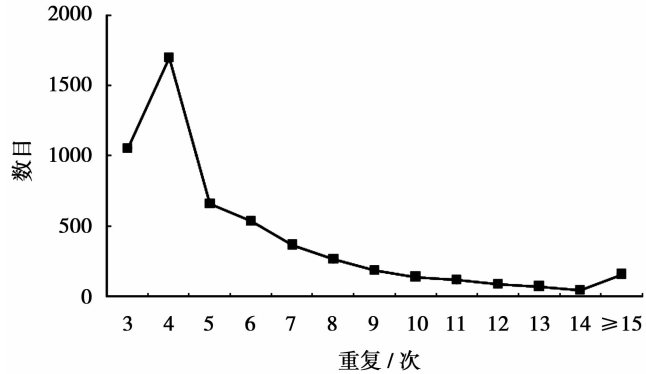


图 1 SSRs 重复次数分布

表 2 芭蕉属主要重复基元的分布

<i>M. acuminata</i>			<i>M. paradisiaca</i>			<i>M. balbisiana</i>		
基元	数量	比例/%	基元	数量	比例/%	基元	数量	比例/%
AG/CT	410	76.35	AG/CT	373	87.97	AG/CT	189	85.91
AC/GT	42	7.82	AC/GT	23	5.42	AC/GT	9	4.09
AT/AT	84	15.64	AT/AT	28	6.61	AT/AT	20	9.09
CG/CG	1	0.19	CG/CG	0	0.00	CG/CG	2	0.91
合计	537	100.00	合计	424	100.00	合计	220	100.00
AAC/GTT	44	2.94	AAC/GTT	16	1.91	AAC/GTT	8	1.57
AAG/CTT	317	21.19	AAG/CTT	247	29.44	AAG/CTT	123	24.07
AAT/ATT	35	2.34	AAT/ATT	6	0.72	AAT/ATT	4	0.78
ACC/GGT	106	7.09	ACC/GGT	67	7.99	ACC/GGT	49	9.59
ACG/CGT	74	4.95	ACG/CGT	24	2.86	ACG/CGT	20	3.91
ACT/AGT	46	3.07	ACT/AGT	29	3.45	ACT/AGT	16	3.13
AGC/GCT	198	13.24	AGC/GCT	108	12.87	AGC/GCT	66	12.92
AGG/CCT	335	22.39	AGG/CCT	201	23.96	AGG/CCT	171	33.46
ATC/GAT	110	7.35	ATC/GAT	73	8.70	ATC/GAT	43	8.42
CCG/CGG	231	15.44	CCG/CGG	68	8.10	CCG/CGG	11	2.15
合计	1 496	100.00	合计	839	100.00	合计	511	100.00
AAAC/GTTT	18	4.60	AAAG/CTTT	48	22.54	AAAC/GTTT	8	4.35
AAAG/CTTT	48	12.28	AAAT/ATTT	13	6.10	AAAG/CTTT	16	8.70
AAAT/ATTT	18	4.60	AAGG/CCTT	21	9.86	AAAT/ATTT	32	17.39
AAGG/CCTT	25	6.39	AAGC/GCTT	16	7.51	AAGG/CCTT	9	4.89
AGGG/CCCT	22	5.63	CCAT/ATGG	11	5.16	AGGT/ACCT	19	10.33
ACCC/GGGT	14	3.58	AGGG/CCCT	12	5.64	AGGG/CCCT	9	4.89
NNNN	246	62.92	NNNN	92	43.19	NNNN	91	49.45
合计	391	100.00	合计	213	100.00	合计	184	100.00
AAAAG/CTTTT	12	10.62	AAGAG/CTCTT	17	28.33	AAAAT/ATTTT	4	11.11
NNNNN	101	89.38	NNNNN	43	71.67	NNNNN	32	88.89
合计	113	100.00	合计	60	100.00	合计	36	100.00
NNNNNN	173	100.00	NNNNNN	109	100.00	NNNNNN	110	100.00
总计	2710		总计	1 645		总计	1 061	

2.2 芭蕉属 3 个种 EST-SSR 的特点

在 3 个种 *M. acuminata*, *M. paradisiaca* 和 *M. balbisiana* 中, 分别搜索到 SSRs 为 2 710、1 645 和 1 061 条, 其二核苷酸重复基元均是以 AG/CT 为主导重复类型, 分别占 76.35%、87.97% 和 85.91%; 三核苷酸重复基元以 AAG/CTT 与 AGG/CCT 占优势, 分别占三核苷酸 SSRs 的 43.58%、53.40% 和 57.53%; 四核苷酸优势重复基元在 *M. acuminata* 和 *M. paradisiaca* 中是 AAAG/CTTT (分别 12.28% 和 22.54%), 而 *M. balbisiana* 则是以 AAAT/ATTT 为主 (17.39%); 其它二、三、四、五和六核苷酸重复基元类型出现频率较低, 分布也相对分散 (表 2)。

2.3 地涌金莲 EST-SSR 多态性引物筛选

在包含有 SSR 的 EST 序列中, 随机挑选 238

对进行 SSR 引物设计以开发地涌金莲多态性 SSR 引物。共有 116 对 (48.74%) 引物有扩增产物, 78 对扩增出条带清晰的目的片段, 有效扩增率为 32.77%, 49 对引物显示出多态性, 多态率为 64.10%。用 15 对多态性引物对来自 4 个地涌金莲野生种群的 24 个个体进行多态性检测 (图 2), 扩增出的等位基因数范围是 2~7 个, 平均 3.067 个, 其表观杂合度 (H_o) 范围是 0.042~0.750, 平均 0.250; 期望杂合度 (H_e) 范围是 0.232~0.823, 平均 0.522。这些引物中, 有 8 对在 $P < 0.01$ 时偏离 HWE 平衡, 占所检测多态性引物的 53.33% (表 3)。

表 3 地涌金莲多态性 EST-SSR 引物信息

位点	重复基元	引物序列 (5'-3')	等位基因/个	退火温度/°C	产物大小/bp	H_o	H_e	HWE
smu27	(AGCTGA) ₄	F:GGCAGTCATCGCGTCTCTA R:AAAGTTCTGAGGCCGAGCG	2	56	238~250	0.261	0.232	0.515
smu150	(AGATCA) ₄	F:GCAAAGCAGCAGCAGCAAGC R:TCGATCATGGCGCTCCCAAG	2	62	202~208	0.000	0.474	0.000*
smu224	(GGA) ₅	F:CAGCCTGTGGCGATCGGTAT R:GCCCTTGACGGTGAGACAGG	2	55	244~247	0.000	0.383	0.000*
smu225	(AAG) ₅	F:CACCTGCCTCCCTTCCTTGC R:TGCCTTCCTTGCCGAGACAC	2	59	182~187	0.042	0.361	0.000*
smu233	(GCC) ₇	F:AGCTCGACAAGGACCGCAAG R:ACAAGCGCAACCGCATAACG	2	62	187~190	0.000	0.474	0.000*
cmu24	(TCT) ₅	F:ACCTGGGGCTGTTCCTTTGC R:TGCCCTTCTCATCTGTGGTATGC	4	59	250~262	0.130	0.714	0.000*
cmu35	(AAG) ₁₃	F:ACAAGATGAGGCCACTGCGG R:ACAAACGCGAACCCCTTGCT	3	61	277~286	0.000	0.557	0.000*
cmu52	(GTG) ₄	F:GCGGTGGTATTCCCCACGAG R:AAGCCCAGATCTTGCCTGC	2	51	235~238	0.250	0.383	0.078
cms3	(CT) ₁₀	F:TGTTGGTAGCTGGAATCCGTC R:TCGATTCTTCCTTCAACTTCCA	4	53	166~178	0.522	0.743	0.015
cms4	(CT) ₁₉	F:AGGGCACACCAAGGTGCAAT R:ACAAGCAAGCAAAAAGCGACGA	3	56	156~164	0.667	0.547	0.095
mms7	(GATACT) ₃ (CAG) ₄	F:AAGAGCGAAACGGGTGGTGA R:TCCGGGGAATGACATTGTCTGA	2	56	212~215	0.000	0.454	0.000*
mms14	(TCA) ₅ (TTC) ₆	F:CTGGCAGCTCTTCGGCGTAT R:GACCGCGCGCTCATTAACA	7	57	232~262	0.333	0.823	0.000*
mms25	(TC) ₁₉ (TG) ₁₃	F:GCAGCCGAAAAAGGTGGTCC R:GCCGAAGACCCACTGGTTGA	5	60	136~160	0.750	0.743	0.273
musf32	(CT) ₁₃	F:TCGGACGTGTTGTCTGTCGG R:ACAGCGTCCGGACCTCGATA	2	57	266~274	0.130	0.264	0.010
musf43	(CA) ₁₃	F:CAACCGACAGGTGGCTTGT R:CGTCATGGCTACCTCCAGGCT	4	62	197~207	0.667	0.679	0.043
平均			3.067			0.250	0.522	

注: H_o :表观杂合度; H_e :期望杂合度; * 偏离哈迪-温伯格平衡 ($P < 0.01$)。

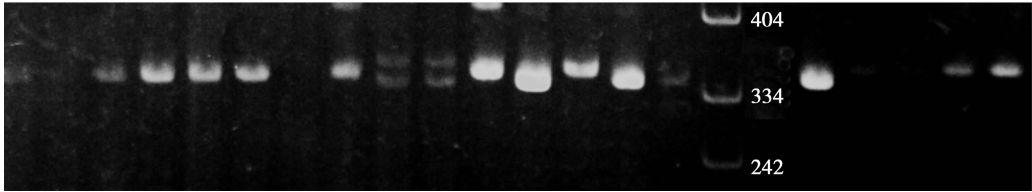


图2 引物 cmu24 扩增带谱

3 讨论

在芭蕉属植物 21 129 条无冗余 EST 序列中,检测到 SSR 位点 5 416 个,检出率是 25.63%,平均每隔 2.49 kb 就有 1 个 SSR,检出率远高于小麦(3.2%)、水稻(4.7%)、玉米(1.5%)、高粱(*Sorghum* spp., 3.6%)和甘蔗(*Saccharum* spp., 2.9%)^[3,17-19],同时也高于在香蕉(*Musa nana* Lour.) EST 序列中 SSR(5.3%)的检出率^[20]。这些差异可能是物种间真实的 SSR 差异或各物种在 GenBank 中可用来分析的 EST 数量不同,也可能是搜寻 SSR 时所用软件或所设定最低重复数不同所致。Varshney 等^[21]认为,高的 EST-SSRs 检出率可能与该物种基因组较小有关。研究表明,三核苷酸在多数植物中分布几率较高^[21]。在芭蕉属植物中,三核苷酸重复序列占 SSRs 的 52.55%,处于主导地位,这与前人对多种植物 EST-SSRs 分析结果一致^[17,22-23]。

当 4 种不同的碱基随机组合时,若 EST-SSRs 数目足够大且无偏倚性,将可能分别产生 2、4、10、33、102 和 350 种单、二、三、四、五、六核苷酸重复^[24]。本研究显示,在芭蕉属 3 个物种 *M. acuminata*, *M. paradisiaca* 和 *M. balbisiana* 中,二、三和四核苷酸重复基元均表现出明显的偏倚性,AAG/CTT 与 AGG/CCT 是三核苷酸重复的主要类型,分别占 SSRs 的 43.58%、53.40% 和 57.53%,这与王静毅^[20]对香蕉 EST-SSRs 分析结果略有差异。此外,GGC/CCG 和 ACG/TCG 在水稻、玉米、高粱和大麦(*Hordeum vulgare* L.) 中是主要三核苷酸重复类型,而小麦中是 AAC/TTG^[17]。产生这种差异的原因除所用 EST 数目不同或 SSR 最小长度设定不同之外,也可能是由于不同物种自身 EST-SSR 的特点及不同研究所选用的 EST 来源有差异所致。与多数植物相似^[17-18, 25],芭蕉属 3 个种的二核苷酸重复基元均是以 AG/CT 为主导重复类型,而 CG/CG 重复数量最少。

EST-SSR 来自编码区,侧翼序列在物种之间高度保守,在物种间具有良好的通用性^[3-5]。本研究中,随机挑选并设计的 238 对芭蕉属 EST-SSR 引物

中,共有 116 对引物在地涌金莲中有扩增产物,其中 78 对可以扩增出清晰、稳定的目的产物,占合成引物的 32.77%,处于中等水平^[26]。160 对引物无扩增产物或引物扩增产物与预期目的片段大小不符,可能是因为对应 DNA 扩增区间有内含子导致产物大于预期片段大小,或是引物位于转座子或重复 DNA 区域内而产生偏离目的片段的产物,也可能是由于该位点在不同属的植物间具有较大的变异,这也印证了 EST-SSR 具有序列保守性。经初步检测,49 对具有多态性,占有扩增产物引物的 64.10%。每对多态性引物检测出的等位基因为 2~7 个,平均 3.067 个,低于地涌金莲基因组 SSR 引物^[13]。研究表明,通常 EST-SSR 基元重复次数较 genomic-SSR 少,因此,其多态性也低^[3,27-28],此外,基因编码区序列高度保守也是 EST-SSR 多态性较低的原因之一。

参考文献:

- [1] Zane L, Bargelloni L, Patarnello T. Strategies for microsatellite isolation: a review [J]. *Molecular Ecology*, 2002,11(1):1-16
- [2] Squirrel J, Hollingsworth P M, Woodhead M, et al. How much effort is required to isolate nuclear microsatellites from plants? [J]. *Molecular Ecology*, 2003,12:1339-1348
- [3] Cordeiro G M, Casu R, McIntyre C L, et al. Microsatellite markers from sugarcane (*Saccharum* spp.) ESTs cross transferable to erianthus and sorghum [J]. *Plant Science*, 2001,160(6):1115-1123
- [4] Kuleung C, Baenziger P S, Dweikat I. Transferability of SSR markers among wheat, rye and triticale [J]. *Theoretical and Applied Genetics*, 2004,108(6):1147-1150
- [5] 郑丽珊,石玉真,王静毅,等. 棉花 EST-SSRs 在香蕉中的通用性 [J]. *中国农学通报*, 2008,24(1):33-37
- [6] Falke K C, Susic Z, Hackauf B, et al. Establishment of introgression libraries in hybrid rye (*Secale cereale* L.) from an Iranian primitive accession as a new tool for rye breeding and genomics [J]. *Theoretical and Applied Genetics*, 2008,117:641-652
- [7] Tahan O, Geng Y P, Zeng L Y, et al. Assessment of genetic diversity and population structure of Chinese wild almond, *Amygdalus nana*, using EST- and genomic SSRs [J]. *Biochemical Systematics and Ecology*, 2009,37(3):146-153
- [8] 徐小彪,姜春芽,廖 娇,等. 中华猕猴桃矮型性状 EST-SSR 连锁

- 标记的筛选[J]. 园艺学报, 2010, 37(4): 553-558
- [9] Thiel T, Michalek W, Varshney R K, *et al.* Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.) [J]. Theoretical and Applied Genetics, 2003, 106(3): 411-422
- [10] 吴征镒. 云南植物志(第二卷)[M]. 北京: 科学出版社, 1979
- [11] Ma H, Pan Q J, Wang L, *et al.* *Musella lasiocarpa* var. *rubibracteata* (Musaceae), a New Variety from Sichuan, China [J]. Novon, 2011, 21: 349-353
- [12] 潘庆杰, 李正红, 王 雁, 等. 地涌金莲野生与栽培种群遗传多样性的 RAPD 分析[J]. 林业科学研究, 2007, 20(5): 668-672
- [13] Yang C Y, Huang Y, Long C L. Isolation and characterization of 17 polymorphic microsatellite loci for *Musella lasiocarpa* (Musaceae) [J]. Hortscience, 2009, 44(7): 2041-2042
- [14] Doyle J J, Doyle J L. Isolation of plant DNA from fresh tissue [J]. Focus, 1990, 12: 13-15
- [15] Huang X, Madan A. CAP3: A DNA sequence assembly program [J]. Genome Research, 1999, 9: 868-877
- [16] Rozen S, Skaletsky H J. PRIMER 3 on the www for general users and for biologist programmers [C] // Krawetz S, Misener S. Methods in Molecular Biology. Totowa, New Jersey, Humana Press, 2000: 365-386
- [17] Kantety R V, Rota M L, Matthews D E, *et al.* Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat [J]. Plant Molecular Biology, 2002, 48: 501-510
- [18] Nicot N, Chiquet V, Gandon B, *et al.* Study of simple sequence repeat (SSR) markers from wheat expressed sequences tags (ESTs) [J]. Theoretical and Applied Genetics, 2004, 109: 800-805
- [19] Cardle L, Ramsay L, Milbourne D, *et al.* Computational and experimental characterization of physically clustered simple sequence repeats in plants [J]. Genetics, 2000, 156: 847-854
- [20] 王静毅, 陈业渊, 刘伟良, 等. 香蕉 EST-SSRs 标记的开发与应用 [J]. 遗传, 2008, 30(7): 933-940
- [21] Varshney R K, Thiel T, Stein N, *et al.* In silico analysis on frequency and distribution of microsatellites in ESTs of some cereal species [J]. Cellular and Molecular Biology Letters, 2002, 7(2A): 537-546
- [22] Chapman M A, Hvala J, Strever J, *et al.* Development, polymorphism and cross-taxon utility of EST-SSR markers from safflower (*Carthamus tinctorius* L.) [J]. Theoretical and Applied Genetics, 2009, 120: 85-91
- [23] Huang H, Lu J, Ren Z B, *et al.* Mining and validating grape (*Vitis* L.) EST-SSR markers for genotyping and mapping [J]. Molecular Breeding, 2010, DOI: 10.1007/s11032-010-9477-2
- [24] Rota L R, Kantety R V, Yu J K. Nonrandom distribution and frequencies of genomic and EST-derived microsatellite markers in rice, wheat and barley [J]. BMC Genomics, 2005, 6: 23
- [25] Feng S P, Li W G, Huang H S, *et al.* Development, characterization and cross-species/genera transferability of EST-SSR markers for rubber tree (*Hevea brasiliensis*) [J]. Molecular Breeding, 2009, 23: 85-97
- [26] Ellis J R and Burke J M. EST-SSRs as a resource for population genetics analysis [J]. Heredity, 2007, 99(2): 125-132
- [27] Areshchenkova T, Ganai M W. Comparative analysis of polymorphism and chromosomal location of tomato microsatellite markers isolated from different sources [J]. Theoretical and Applied Genetics, 2002, 104: 229-235
- [28] Ramsay L, Macaulay M, Ivanisovich S, *et al.* A simple sequence repeat-based linkage map of barley [J]. Genetics, 2000, 156: 1997-2005